DOCUMENT RESUME

ED 122 255 CS 002 617

AUTHOR Tucker, Elizabeth Sulzby
TITLE Instructional Level in Reading; Including a Critique of the Informal Reading Inventory.
PUB DATE 69
NOTE 85p.; Paper presented at the Annual Meeting of the Virginia State Reading Association (Richmond, virginia, March 6-8, 1975)

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS *Academic Achievement; Achievement Tests; Developmental Reading; Elementary Education; Individualized Reading; *Informal Reading Inventory; *Reading Achievement; Reading Diagnosis; *Reading Instruction; Reading Research; *Reading Tests

ABSTRACT

The purpose of this study was to describe selected tests which have been used to yield instructional levels in reading; to examine the bases on which such tests have been designed in light of their relevance to instructional level; and to inspect the relation of test standards to selected methods of teaching reading in the elementary school. The first two sections of the document discuss such topics as the importance of the study, definitions of terms used, reading achievement tests, bases on which tests were constructed, and relation of standards to methods of teaching reading. The third section is a summary of the study and a delineation of the conclusions, including the conclusion that reading tests were found to be based on three types of standards: proficiency standards, normative populations, and analyses of reading skills. The informal reading inventory was found to be the only generally accepted test that would accurately yield a child's instructional level. (TS)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED122255

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

INSTRUCTIONAL LEVEL IN READING

including a critique

of the Informal Reading Inventory

Elizabeth Sulzby Tucker
July 1969

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

Elizabeth Sulzby Tucker

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

CS 002 617

TABLE OF CONTENTS

CHAPTER		PAGE
I.	THE PROBLEM AND DEFINITIONS OF TERMS USED . . .	1
	The Problem	1
	Statement of the problem	1
	Importance of the study	1
	Definitions of Terms Used	6
	Instructional level	6
	Basal reader	6
	Developmental reading	6
	Individualized reading	7
II.	READING ACHIEVEMENT	8
	Tests	8
	Basal reader proficiency tests	8
	Standardized achievement tests	13
	Diagnostic tests	25
	Informal reading inventories	37
	Bases Upon Which Tests were Constructed . . .	55
	Proficiency standards	55
	Norming populations	57
	Analyses of reading skills	59
	Relation of Standards to Methods of Teaching Reading	61

CHAPTER PAGE

"Whole-word" basal readers 61

Phonic readers 63

Individualized reading 65

III. SUMMARY AND CONCLUSIONS 69

Summary 69

Conclusions 73

BIBLIOGRAPHY 75

CHAPTER I

THE PROBLEM AND DEFINITIONS OF TERMS USED

Placing a child in reading material of appropriate difficulty and teaching him at a rate most effective to learning has been a procedure central to the theory that reading is developmental. Teachers of remedial reading have made the practice of determining the instructional levels of pupils the most important part of diagnosis. Testing methods have been devised that claim to enable the teacher to choose material of proper difficulty for the student.

I. THE PROBLEM

Statement of the problem. It was the purpose of this study (1) to describe selected tests which have been used to yield instructional levels in reading, (2) to examine the bases upon which such tests have been designed in light of their relevance to instructional level, and (3) to inspect the relation of test standards to selected methods of teaching reading in the elementary school.

Importance of the study. Placing a child in material which is too difficult has been blamed for retarding the

reading progress of many children in American schools.[1] Harris cited a study by Gertrude Whipple in which "use of inappropriate reading material" was the most frequently mentioned cause of reading difficulty by experienced teachers.[2] In one investigation conducted by Jeanne Chall for the U. S. Office of Education First Grade Reading Studies, "appropriateness of the level of difficulty of the reading lessons" was one teacher characteristic which related significantly with reading achievement after one year of instruction. While the authors noted several limitations of the study, the results agreed with the observations of Bond, Tinker, Harris, and Betts.[3]

Emerald Dechant declared that "about ten to fifteen million children in regular classrooms attempt to read books that are too difficult for them and that this

---

[1]Guy L. Bond and Miles A. Tinker, Reading Difficulties: Their Diagnosis and Correction (second edition; New York: Appleton-Century-Crofts, 1967), p. 139; Emmett Albert Betts, Foundations of Reading Instruction (New York: American Book Company, 1946), p. 452; and Arthur W. Heilman, Principles and Practices of Teaching Reading (second edition; Columbus, Ohio: Charles E. Merrill Publishing Company, 1961), p. 8.

[2]Albert J. Harris, How to Increase Reading Ability (fourth edition, revised; New York: David McKay Company, Inc., 1961), p. 16.

[3]Jeanne Chall and Shirley Feldmann, "First Grade Reading: An Analysis of the Interactions of Professed Methods, Teacher Implementation and Child Background," The Reading Teacher, XIX (May, 1966), 572.

interferes with reading progress."[4] Madeline Hunter, in 1964, reviewed the literature related to difficulty level. She stated, as a "less widely known premise," that:

> Materials that are inappropriate or improperly or partially learned are not only wasteful in themselves but they interfere with previous and future learning. Psychological research vividly demonstrates this retroactive and proactive interference.[5]

Albert J. Harris outlined objectives for reading instruction that emphasized individual differences in children and the responsibility of the teacher in providing for those differences.[6] In an article directed toward the general public, Samuel G. Sava asserted:

> The key to excellent instruction at any level is tailoring it to each student's ability, so that his academic work neither frustrates him with too much difficulty, nor bores him with repetition.[7]

Emphasis upon child development and individual differences led educators to delve further into the developmental nature of learning and, in particular,

---

[4]Emerald Dechant, Diagnosis and Remediation of Reading Disability (West Nyack, New York: Parker Publishing Company, 1968), p. 35.

[5]Madeline Hunter, "The Dimensions of Ungrading," The Elementary School Journal, LXV (October, 1964), 24.

[6]Harris, loc. cit.

[7]Samuel G. Sava, "When Learning Comes Easy," Saturday Review, LI (November, 1968), 119.

of learning to read.[8] Individual differences in the growth and development of children made evident the impossibility and the undesirability of expecting all children in a classroom to proceed at a similar rate.[9]

The problem of meeting the needs of children based upon their individual differences in abilities and progress was attacked by plans for grouping. Nila Banton Smith reviewed numerous plans that have been used in America since 1870.[10] No matter what organizational plan for grouping was followed, Smith found that "how to choose the right material for the individual child posed problems."[11]

Psychologists had laid the foundation for a theory declaring that each child should be taught at an appropriate level of difficulty. E. L. Thorndike's opinion about the desirable amount of unknown material in supplementary books was cited by Betts in his Foundations of Reading Instruction as a basis for the reading standards of his informal reading inventory. Thorndike had claimed: "Words unknown

---

[8]Frances L. Ilg, "The Child from Three to Eight, with Implications for Reading," Issues and Innovations in the Teaching of Reading, Joe Frost, editor (Glenville, Illinois: Scott, Foresman and Company, 1967), pp. 81-8.

[9]Bond and Tinker, op. cit., pp. 45-6.

[10]Nila Banton Smith, Reading Instruction for Today's Children (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963), pp. 108-26.

[11]Ibid., p. 139.

to the reader should occur only rarely. A reasonable standard is not over 1 in 200."[12]

Teaching machines, such as those constructed by Sidney Pressey and B. F. Skinner, were built to "insure 90 to 95 per cent correct student responses to a frame."[13] Such norms implied that a student could tolerate only 5 to 10 per cent of unknown material. Skinner believed that the programmer's responsibility was to construct a program that would "be arranged in a plausible developmental order."[14] This arrangement, he felt, would help insure a 90 to 95 per cent successful learning experience.

Research into individual differences, child development, and learning load led reading authorities and textbook writers into the study of readability, or the difficulty level of reading materials.[15] At the same time, tests were constructed to aid teachers in determining at what reading levels children were performing.[16]

---

[12]Betts, op. cit., p. 447.

[13]M. Ray Loree, Psychology of Education (New York: The Ronald Press Company, 1965), p. 229.

[14]B. F. Skinner, "Teaching Machines," Readings for an Introduction to Psychology, Richard A. King, editor (New York: McGraw-Hill Book Company, Inc., 1961), p. 176.

[15]Gerald A. Yoakam, Basal Reading Instruction (New York: McGraw-Hill Book Company, Inc., 1955), pp. 20-21.

[16]Smith, op. cit., pp. 110-111.

This study was designed to examine the testing means available to the teacher to choose the correct material for the individual child. A further attempt was made to outline various limitations and advantages of each method when applied to selected methods of teaching reading in the elementary school.

II. DEFINITIONS OF TERMS USED

Instructional level. In this paper instructional level was defined as the level of book which the pupil reads with success in word recognition and comprehension, yet which contains a degree of unlearned material (i.e., vocabulary, concepts, and skills). The amount which the child could learn without frustration was discussed from the viewpoints of several authorities.

Basal reader. Basal reader was used to mean a series of books designed for sequential reading instruction. The term included the readers (called "whole-word" basal readers herein) which teach a body of sight words before instructing the student in word attack skills, such as phonics and structural analysis. Basal readers also implied phonic readers or readers which begin with phonic elements, then present the child with stories and exercises emphasizing phonetically regular words.

Developmental reading. Developmental reading was

to mean sequentially taught reading skills based upon growth patterns of children. The growth patterns included were mental growth, physical growth, perceptual growth, experiential growth, and emotional growth.[17]

Individualized reading. Personalized reading and individualized reading were placed into one category. Individualized reading was defined as a system of reading in which the child selects his own reading materials from which the teacher chooses reading skills to teach, either to the individual child alone, or to small, temporarily-formed groups. Systems which used basal readers in addition to library books were also included.

---

[17]Smith, op. cit., pp. 26-50.

CHAPTER II

READING ACHIEVEMENT

Between 1910 and 1920 reading tests were developed and used for the first time. Concurrently, intelligence tests were also being used initially in America on a wider scale. Nila Banton Smith attributed a sudden awareness of individual differences in reading progress and potential to the development and use of these two types of tests.[18] Four types of reading tests were selected for discussion in this paper: Basal reader proficiency tests, standardized achievement tests, diagnostic tests, and informal reading inventories. These tests were based upon three important bases: proficiency standards, norming populations, and analyses of reading skills. The tests chosen and their standards were related to two types of basal reader programs and to individualized reading in an attempt to examine the means of determining a child's instructional level which are presently available to teachers.

I. TESTS

Basal reader proficiency tests. Editors of many basal reader series have written proficiency tests to

---

[18]Smith, op. cit., p. 130.

accompany the textbooks. Such tests have been designed to judge the student's progress in learning the vocabulary, skills, and content of the series.[19]

Emmett A. Betts and Carolyn M. Welch composed the Reading-Study Achievement Tests to accompany the Betts Basic Readers by the American Book Company.[20] A separate test was written at each reader level with twelve subtests whose topics varied to fit the reader level.

In 1963 the authors enumerated the following purposes for the Reading-Study Achievement Tests:

1. To motivate pupils through awareness of progress.
2. To identify specific needs basic to successful learning at the next reader level.
3. To provide objective evidence for conferences with administrators, supervisors, and parents.
4. To evaluate areas in which the pupil might require additional instruction before the next level.
5. To obtain objective data on each pupil's record of achievement for use by succeeding teachers.[21]

In 1964 Betts and Welch answered a series of questions concerning these tests, including a question about the purposes for the tests. Rather than re-listing the five

---

[19]Helen M. Robinson, Marion Monroe, and A. Sterl Artley, Guidebook, More Friends Old and New (Chicago: Scott, Foresman and Company, 1963), p. 263.

[20]Emmett A. Betts and Carolyn M. Welch, Overview for Teachers, Reading-Study Achievement Tests (New York: American Book Company, 1964), p. 1.

[21]Emmett A. Betts and Carolyn M. Welch, Manual for Teachers, First-Reader Level, Reading-Study Achievement Tests (New York: American Book Company, 1963), p. 1.

purposes from the 1963 Manual, they summarized the purposes into two categories: diagnosis and measurement of pupil readiness for the next reader level.[22]

Four readiness categories were defined from the scores on the subtests. For each category, the authors indicated a level of basal reader for which the pupil who attained this performance might be ready. Low achievers were said to need help in the same reader level as the test. Average achievers were described as being ready for the next book in the series. High achievers were the pupils perhaps ready to skip one reader level. Teachers were advised to consider having superior achievers skip one or more readers in the Betts series.[23]

Betts and Welch called the Reading-Study Achievement Tests standard tests, rather than standardized tests. They defined standard tests as tests which "measure mastery of learnings," rather than compare one pupil's achievement with that of other students in a grade; these tests were written to measure "specific skills and abilities taught at each reader level."[24]

---

[22]Betts and Welch, Overview, p. 1.

[23]Betts and Welch, Manual, First-Reader Level, pp. 1-2.

[24]Betts and Welch, Overview, p. 1.

Each reader level test included twelve subtests (thirteen at 3-2 and 4th reader level). The number of items varied from 130 at pre-primer level to 440 at 4th reader.[25] In the Overview for Teachers, the authors outlined the score standards in terms of percentages and reader characteristics:

> In general, a superior achiever should attain a score of 100% on the subtests of phonic skills and thinking abilities. High achievers should miss no more than one item on a subtest. Average achievers generally should attain scores of 95% to 97% on these subtests. In any event, as indicated in the Manual for Teachers, the pupils should show signs of success and no signs of difficulty.
>
> Below average achievers will usually show symptoms of difficulty . . . .[26]

Betts and Welch recommended that all subtests be given, unless local school policy or teacher conviction rules against the teaching of certain "word perception skills." They recommended that sufficient time be allowed for the test, with a varied order for administering the subtests, if needed to facilitate a child's reading performance.[27]

The Reading-Study Achievement Tests were called

---

[25] Ibid., pp. 3-4.

[26] Ibid., p. 2.

[27] Ibid., pp. 1-2.

standard tests, based upon the belief that the skills tested should be mastered before the pupil progresses to a higher level reader.[28] A similar purpose was established by other basal reader proficiency tests, such as the Basic Reading Test to the Curriculum Foundation Series by Scott, Foresman and Company.[29] The American and Scott, Foresman authors claimed that the tests indicated instructional levels in the specific skills and content taught in the specific series for which they were designed; they did not claim that instructional level scores from their tests could be generalized to other reader series. Betts pointed out the differences in content and vocabulary load in different basal reader series.[30] Research in readability in recent years emphasized that such differences continue to exist, especially between reader series with greatly different methods of beginning instruction.[31]

Authors of basal reader proficiency tests also cautioned that other means of evaluation should be used

---

[28] Ibid., p. 1.

[29] Robinson et. al., loc. cit.

[30] Betts, Foundations of Reading Instruction, p. 475.

[31] E. B. Coleman, "Experimental Studies of Readability," Elementary English, XLV (February, 1968), 166-78; and George R. Clare, "The Role of Word Frequency in Readability," Elementary English, XLV (January, 1968), 21-2.

along with the proficiency tests. The Basic Reading Test by Scott, Foresman and Company included checklists to be used by the teacher for an informal evaluation in word perception, interpretation, and attitudes and habits.[32] The Betts Basic Readers by American Book Company advised that their tests could be complemented by teacher observation, unit tests, and an informal inventory.[33] While the proficiency tests were designed to be given at the end of instruction in a textbook, these other measures were recommended to be used informally at any time. This flexibility of testing was cited by Smith as being indicative of recent trends in reading instruction and was advised by Gray and Betts. Smith emphasized Betts' conviction that a child should not be required to sit in a group when his ability has outpaced or fallen behind the others, lest he develop faulty reading habits which will have to be broken.[34]

Standardized achievement tests. Standardized achievement tests were designed without allegiance to any specific set of readers. They were prepared to inform

---

[32]Robinson et. al., op. cit., pp. 12-16.

[33]Betts and Welch, Manual, Reading-Study Achievement Tests, p. 1.

[34]Nila Banton Smith, American Reading Instruction (second edition; Newark, Delaware: International Reading Association, 1965), pp. 293-94.

a teacher how well a student can read certain passages in comparison with his age and grade peers.[35] Edward B. Fry listed four legitimate purposes for which a teacher or school system could use a reading achievement test: (1) to compare a student's achievement with the achievement of other students (the standardization group); (2) to determine the degree to which the reading curriculum is being learned; (3) to compare a pupil's achievement with that of other students in a class, or to compare the achievement of instructional groups in the class; and (4) to diagnose individual or group strengths and weaknesses. Fry expanded item (2) by an illustration: "Teachers might want an answer to such a legitimate question as 'Is Johnny ready to start using the second grade book?'"[36] According to Fry, determination of instructional level was one of the purposes of standardized reading tests.

An example of a widely-used standardized achievement test was the series written by Arthur I. Gates, last revised in 1957. The test battery was designed for reading levels from grade one through grade ten. The Gates Primary Reading Tests had three subtests for

---

[35]John J. DeBoer and Martha Dallman, The Teaching of Reading (revised edition; New York: Holt, Rinehart and Winston, Inc., 1965), p. 314.

[36]Edward B. Fry, "Developmental Reading Tests," The Sixth Mental Measurements Yearbook, Oscar Krisen Buros, editor (Highland Park, New Jersey: The Gryphon Press, 1965), pp. 1055-56.

grade one and the first half of grade two: Word Recognition, Sentence Reading, and Paragraph Reading. The Gates Advanced Primary Reading Tests, for the second half of grade two and grade three, excluded Sentence Reading as a subtest. For grades three through eight, the Gates Basic Reading Tests provided five subtests: Reading to Appreciate General Significance, Reading to Understand Precise Directions, Reading to Note Details, Reading Vocabulary, and Level of Comprehension. The Gates Reading Survey was devised as a shorter, more economical version of the Gates Basic Reading Test, but the new test extended through grade ten, with only three subtests: Speed and Accuracy, Reading Vocabulary, and Level of Comprehension.[37]

Albert J. Harris prefaced his description of the Gates test with these remarks:

> The major purpose of a survey test reading [sic] is to give a fairly accurate measure of the level of difficulty at which a pupil can read. These tests generally have long time limits, so that the pupil is stopped primarily by the increasing difficulty of the items rather than by the calling of time.[38]

Harris called the Gates tests good examples of such tests, but he criticized them for giving unrealistic scores for poor readers, "who often make artificially high scores

---

[37]Arthur I. Gates, Manual for the Gates Advanced Primary Reading Tests (New York: Teachers College Press, 1961), pp. 11-12.

[38]Harris, How to Increase Reading Ability, pp. 171-72.

on the basis of successful guessing despite a penalty for wrong answers."[39]

Kenneth D. Hopkins made this evaluation of the Gates tests:

> Reliability is excellent, although the standardization and norming leave some unanswered questions which limit generalizability and interpretation. More information regarding content validity would have been desirable yet the tests appear to be useful survey measures.[40]

The Gates tests were replaced in 1965 by the Gates-MacGinitie Reading Tests, by Arthur I. Gates and Walter H. MacGinitie, "based upon a new, nationwide standardization."[41] The Technical Manual was created to provide the added information referred to by Hopkins: item analysis, establishment of norms, standardization, correlation between scores on the subtests, and correlation with IQ as obtained from the Lorge-Thorndike Intelligence Tests.[42] A Technical Manual was furnished with every test packet in addition to the Teacher's Manual.

---

[39] Ibid., p. 172.

[40] Kenneth D. Hopkins, "Gates Advanced Primary Reading Tests," The Sixth Mental Measurements Yearbook, p. 1059.

[41] Arthur I. Gates and Walter H. MacGinitie, Technical Manual, Gates-MacGinitie Reading Tests (New York: Teachers College Press, 1965), p. 1.

[42] Ibid., pp. 2-5.

A convenient table on the Gates-MacGinitie Reading Tests was printed by Arthur W. Heilman who listed the test first among his representative tests:

	Grade	Subtests	Forms
Primary A	1	Vocabulary, Comprehension	1-2
Primary B	2	Vocabulary, Comprehension	1-2
Primary C	3	Vocabulary, Comprehension	1-2
Primary CS	2.5-3	Speed, Accuracy	1-2-3
Survey D	4-5-6	Speed, Vocabulary, Comprehension	1-2-3
Survey E	7-8-9	Speed, Vocabulary, Comprehension	1-2-3[43]

(In Survey D and E, Accuracy scores were included within the Speed subtest.[44])

The Vocabulary sections of Primary A and Primary B were composed of forty-eight items consisting of a picture with four words. The pupil was asked to circle the word that "goes best with the picture."[45] Primary C instructed the pupil to choose the word that "goes best" with a picture for the first twelve of fifty-two Vocabulary items, then to choose a synonym or the word that "means most nearly the same as" the word in dark print for the remaining items.[46] Survey D followed the format of asking the student to find the word that "means most nearly

---

[43]Heilman, op. cit., p. 184.

[44]Arthur I. Gates and Walter H. MacGinitie, Teacher's Manual, Gates-MacGinitie Reading Tests, Survey D (New York: Teachers College Press, 1965), p. 1.

[45]Gates and MacGinitie, Manual, Primary B, p. 2.

[46]Gates and MacGinitie, Manual, Primary C, p. 3.

the same as" the test word for fifty items.[47]

The Comprehension subtests in Primary A and Primary B asked the student to read a sentence or small story and to mark a large cross over the one of four pictures that "goes best with the story."[48] Thirty-four such items were included. Primary C was designed to have the pupil read a short story or paragraph, to read two or three questions about the passage, and then to circle the best answer to each of the forty-eight questions.[49] Survey D was formed somewhat differently, with the Comprehension subtest resembling what is known as the cloze procedure. The pupil was asked to read a short paragraph containing two or three numbered blanks. Numbered lines of words were placed beneath the paragraph, with the student instructed to choose the word that "makes the best sense" in the blank with the matching number. Fifty-two blanks were prepared.[50]

Speed and Accuracy scores were made available by asking the pupil to read thirty-six short paragraphs and to answer a short question from each paragraph. A time

---

[47]Gates and MacGinitie, Manual, Survey D, p. 4.

[48]Gates and MacGinitie, Manual, Primary B, p. 4.

[49]Gates and MacGinitie, Manual, Primary C, p. 4.

[50]Gates and MacGinitie, Test Booklet, Survey D, p. 6.

limit of six minutes was set. Two scores were to be obtained from the single subtest, the total number answered (speed) and the total number correct (accuracy).[51]

The Gates-MacGinitie Reading Tests were designed to yield three scores from the raw scores: grade scores, or grade equivalents; percentile ranks; and standard scores. The latter two scores were based on the assumption that the pupil would take the test for his grade placement either in the months of October, February, or May.[52] The standard score was included so that scores for a class might be averaged if the teacher desired; however, the authors discouraged averaging the subtest scores for an individual student, particularly if the subtest scores differed significantly.[53]

Standardized achievement tests, like the Gates-MacGinitie tests, were praised for having long subtests, giving pupils an increased and more reliable difficulty-range within which to reach their top level of silent reading performance. Reading specialists were found to differ in their opinions concerning whether such tests could yield a reliable instructional level for a child.

---

[51]Gates and MacGinitie, Manual, Survey D, p. 10.

[52]Gates and MacGinitie, Manual, Primary C, pp. 7-8.

[53]Gates and MacGinitie, Manual, Survey D, p. 10.

Bond and Tinker recommended the use of a good reading survey to locate disabled readers and to yield "a fairly adequate measure of the grade level at which a pupil can read."[54] They indicated a preference for the old Gates tests as one of two examples of such tests. The format of such reading surveys recommended their use, according to Bond and Tinker, in that they begin with easy items and progress at a relatively even rate of difficulty, over a range of several grades.[55]

Harris stated that only one reading level is gained from such a reading test, however, as compared with informal tests which can give independent, instructional, and frustrational levels for each child. Harris called the reading level obtained from a standardized reading survey test "instructional level rather than the independent level, which is usually lower."[56] Harris was departing from his usual use of the term instructional, for he believed standardized tests overestimated instructional level.[57]

---

[54]Bond and Tinker, op. cit., p. 196.

[55]Ibid.

[56]Harris, How to Increase Reading Ability, p. 180.

[57]Albert J. Harris, Effective Teaching of Reading (New York: David McKay Company, Inc., 1962), 122.

Nila Banton Smith also disputed the claim that standardized tests can give instructional level for the individual child accurately:

> Although standardized reading tests have many valuable uses, rarely do their results pinpoint the grade level at which corrective or remedial instruction should begin.[58]

A more stringent accusation against standardized reading tests for failing to indicate instructional level was written by Morton Botel:

> Most reading specialists have concluded that the grade equivalent score on standardized tests tends to overrate the pupil's reading performance, particularly that of the slow reader or slow learner. Thus, the pupils who are most frequently misplaced are the ones most in need of correct placement.[59]

Harris explained the tendency to misplace poor readers through standardized tests. He asserted that poor readers tend to guess answers, resulting in overestimation of instructional level; that nervous readers may miss simple items, then answer more difficult questions correctly; and that apprehensive readers often give up before reaching the most difficult items that they can read. He stated that standardized reading tests are "less accurate for

---

[58]Nila Banton Smith, Graded Selections for Informal Reading Diagnosis, Grades 4 through 6 (New York: New York University Press, 1963), vi.

[59]Morton Botel, Guide to the Botel Reading Inventory (Chicago: Follett Publishing Company, 1966), p. 5.

poor readers than for good readers."[60]

Support for the views of Smith, Botel, and Harris was furnished by a study at the Reading Clinics of the New York City Board of Education. Standardized reading tests were administered to 1,089 children entering the Reading Clinics. The investigation found that poor readers such as these did tend to mark answers randomly on the admitting test. The authors suggested that these children were accustomed to failure.

After being in the atmosphere of the Reading Clinics, described as reassuring and offering success in reading, each child was given a second form of the same reading test. These children had not yet made reading gains observable in their reading class performance, yet they showed gains on the silent reading test. The items missed were analyzed for type of error; the authors attributed the gains to purposeful marking of answers, rather than the previous random marking.[61]

Random marking would render an unrealistic score, Edward Fry declared when criticizing the Developmental

---

[60]Harris, How to Increase Reading Ability, p. 180.

[61]Stella M. Cohn and Margaretta W. Fite, "Personal-Social Changes Reflected in Reading Accuracy Rating," The Reading Teacher, XVII (November, 1963), 97-99.

Reading Tests, another standardized reading test series, for establishing no cutoff scores at the bottom of the grade equivalent scale. On this particular test, he stated that random marking would yield 1.45 in General Comprehension on a primary form and 4.7 on an intermediate form.[62]

Another aspect of unrealistic marking was attested to by Donald Cleland. He contended that a good reader would often establish a set toward a long standardized silent reading test and would unthinkingly mark answers in response to the test format rather than to individual questions.[63]

Heilman did not indicate a single test alone for establishing instructional level in reading for a pupil.[64] He offered, instead, the following precaution concerning standardized test usage, urging that they should be used to obtain information upon which to build a program of successful instruction:

> Standardized tests are widely used in our schools, and a larger number of them are becoming available each year. Most of these have real merit, yet it is doubtful that reading instruction is improving as a direct result of these tests . . . . The only

---

[62]Fry, op. cit., p. 1056.

[63]Donald L. Cleland, "Clinical Materials for Appraising Disabilities in Reading," The Reading Teacher, XVII (March, 1964), 428-34.

[64]Heilman, op. cit., pp. 193-4.

> justifiable purpose for the use of reading tests is to secure data about an individual's reading ability so that a reading program for the child can be built from the data secured. In actual practice some schools and some teachers gain comfort from the use of tests because they are convinced that testing programs per se have educational value. Testing becomes an end in itself rather than a basis for instruction.[65]

Several suggestions have been offered to improve the use of standardized reading tests in finding instructional level. Wilson suggested that the child should be given a test the level of his reading ability, rather than that of his grade placement if a difference exists.[66] Fry reminded teachers to examine exactly what the test measures, rather than taking subtests at face-value.[67] Traxler's analysis of reading tests has been widely cited by reading specialists cautioning against misuse of tests. Traxler appraised twenty-eight tests and found that they measured forty-nine different skills.[68] Donald Cleland asserted that the best standardized test would only prove as good as the examiner or

---

[65]Heilman, op. cit., p. 181.

[66]Robert M. Wilson, Diagnostic and Remedial Reading (in Merrill's International Education Series, ed. Kimball Wiles. Columbus, Ohio: Charles E. Merrill Publishing Company, 1967), p. 86.

[67]Fry, loc. cit.

[68]Cleland, op. cit., 431.

teacher administering and interpreting it.[69]

Diagnostic tests. Two main types of standardized diagnostic test have been developed, those which test silent reading and those which test oral reading skills. Other tests covered both types of reading. Diagnostic tests were designed to reveal "strong and weak areas in a pupil's repertory of the abilities measured in the subtests."[70] Three tests which were written to reveal pupil strengths and weaknesses were the Gray Oral Reading Tests, the Silent Reading Diagnostic Tests, and the Analysis of Reading Difficulty.

Before his death, William S. Gray designed the Gray Oral Reading Tests along the pattern of the earlier Standardized Oral Reading Paragraphs to accompany silent reading tests in measuring reading growth.[71] Four forms of the test were written in 1954 with the assistance of several experts, taking into consideration a core vocabulary from eleven basal reader series, vocabulary density and syllabic length, sentence length and structural complexity, and topic interest. After construction, the

---

[69]Ibid., 433-34.

[70]Bond and Tinker, op. cit., p. 197.

[71]William S. Gray, Manual to Gray Oral Reading Tests (New York: The Bobbs-Merrill Company, Inc., 1963), p. 3.

passages were tested for equivalence by expert opinion and readability formulae.[72]

Comprehension questions were written for each passage, but they were not submitted to standardization. They were judged by reading specialists not to be for the purpose of diagnosis in specific types of comprehension, but rather "to provide a good measure of the reader's grasp of the meaning of the passage as a whole . . . ."[73] Harris approved of providing no norms for these questions since they were limited to literal comprehension.[74]

The administrator of the Gray test was required to record two types of information as the pupil reads aloud, the time of reading and the errors made.[75] A chart was included to yield a passage score from the number of errors and time (in seconds).[76] The total of these passage scores was then to be located on the grade equivalent charts for each form. Separate charts were provided for boys and girls.[77]

---

[72] Ibid., p. 25.

[73] Harris, The Sixth Mental Measurements Yearbook, p. 1130.

[74] Gray, loc. cit.

[75] Ibid., pp. 3-7.

[76] Ibid., p. 7.

[77] Ibid., pp. 8-15.

The Examiner's Record Booklet included charts beside each passage within which the type of error (aid, gross mispronunciation, partial mispronunciation, omission, insertion, substitution, repetition, inversion) could be recorded. A large summary chart was placed on the front of the record booklet, along with a chart for recording the number of errors, time, passage scores, and comprehension data. A checklist titled "Observations" and a place for comments was also furnished, with blanks for information concerning the pupil's background.[78]

The Manual defined the purpose for the test:

> The first purpose of the test is to assess the level of oral reading achievement. The grade score is usually sufficient to serve this purpose. However, the grade score must be interpreted cautiously for individuals.
>
> All tests are composed of samples of expected performance and, therefore, are subject to error.[79]

The Manual recommended the use of the grade score along with the standard error of measurement "to determine the range within which each pupil's true score probably lies."[80]

---

[78]William S. Gray, Examiner's Record Booklet for the Gray Oral Reading Test, Form B, p. 1.

[79]Gray, Manual, p. 16.

[80]Ibid.

Bliesmer, Harris, and Lohnes praised the new Gray test in reviews in the Sixth Mental Measurements Yearbook, with the exception of the weakness they found in its norms, which are tentative. Lohnes did question the relation of oral reading to silent reading; he stated that he realized the manual also raised the question but did not answer it. He referred to the studies on the subject that have implied that the relationship is complicated by other variables, such as psychological aspects. Only Harris mentioned the utility of the Gray test for obtaining instructional level. He remarked that the norms unduely penalize beginning readers:

> . . . A child could read the preprimer passage perfectly and get partial credit on the primer passage, and still score 1.0, the same as a child unable to recognize a single word.[81]

Some writers contended that oral reading is a good tool in gaining an accurate instructional level. A study by Hardin and Ames found significantly different instructional levels for only nine of twenty-six students when the Gray test was compared to an informal reading inventory; these differences

---

[81]Emery P. Bliesmer, Albert J. Harris, and Paul R. Lohnes, "The Gray Oral Reading Test Tests," The Sixth Mental Measurements Yearbook, pp. 1129-32.

were attributed partially to failure to meet the comprehension standards on the informal reading inventory.[82] Helen Huus summed up advantages of the use of oral reading in judging instructional level:

> Oral reading is used because it is individual, because every word must be included in the reading, and because the reader's grasp of meaning can be deduced in part from his phrasing and intonation.[83]

An example of the second type of diagnostic test was created by Guy L. Bond, Theodore Clymer, and Cyril J. Hoyt. The Silent Reading Diagnostic Tests were designed to be a group test using only silent reading in all subtests. The manual made these claims that relate to determination of instructional level:

> The Silent Reading Diagnostic Tests diagnose silent reading abilities . . . . Research shows that oral and silent reading abilities are not the same. Most reading is silent reading. Therefore, the emphasis in diagnosis should be on silent reading.
>
> The Silent Reading Diagnostic Tests provide a complete diagnosis.[84]

---

[82]Verales B. Hardin and Wilbur S. Ames, "A Comparison of the Results of Two Oral Tests," The Reading Teacher, XXII (January, 1969), 329-34.

[83]Helen Huus, "Basic Reading Skill Instruction in the Total Curriculum," Reading and the Elementary School Curriculum, David L. Shepherd, editor (Newark, Delaware: International Reading Association, 1969), p. 20.

[84]Guy L. Bond, Theodore Clymer, and Cyril J. Hoyt, Manual to Silent Reading Diagnostic Tests (Chicago: Lyons & Carnahan, 1955), p. 2.

The editors asserted that this test, composed of eleven subtests, provides the classroom teacher with an easy-to-administer and complete diagnosis. The subtests (Recognition of Words in Isolation, Recognition of Words in Context, Recognition of Reversible Words in Context, Locating Elements, Syllabication, Locating Root Word, Word Elements, Beginning Sounds, Rhyming Sounds, Letter Sounds, and Word Synthesis) were devised to be administered in two sessions, approximately forty-five minutes per session in whole-class administration.[85]

The test was designed primarily for children of third grade age and above, but the authors described successful experimental use in second grade.[86] The lowest grade equivalent score on the Graphic Profile which is used for scoring appeared to be 1.5, and the highest score, 6.0. Cutout keys were provided for ease of scoring. The first two tests, Words in Isolation and Words in Context, required that each answer be marked either correct, or as an error (classified as Initial, Middle, Ending, or Orientation), and both correct answers and types of errors be recorded along with the scores

---

[85] Ibid., p. 4.

[86] Ibid., p. 2.

from the other subtests. All scores were to be recorded on the Graphic Profile, along with the child's chronological grade, mental grade, grade in school, and reading ability scores as determined by a separate silent reading survey test. The Average Reading grade equivalent, calculated from the silent reading test, was described as the basis for what the authors termed Lines of Importance.[87]

The writers explained that scores above the Lines of Importance indicate a skill area that has been overdeveloped in relation to other skills, and scores below the Lines of Importance indicate a need for additional instruction. "Any measure which extends to or outside of the Lines of Importance is certain to be of diagnostic importance."[88]

The Lines of Importance were to be drawn on either side of the line for Average Reading, extending from a six months range before grade 2.4 to a two year range at grade 5.5 and above.[89] Besides information about unevenness of skill development, and also indication that a child is a candidate for remedial help, no information was given in the manual about determining a child's

---

[87] Ibid., pp. 11-12.

[88] Ibid., p. 2.

[89] Ibid., p. 12.

instructional level.[90]

The manual included one page of technical data covering norming, reliability, and the error keys for tests which require classification. The manual stated:

> In developing the various tests and establishing tentative norms, the Silent Reading Diagnostic Tests were given to all the children in a typical Midwestern community with a population of approximately 20,000, and to other groups from other representative communities.[91]

Emery P. Bliesmer criticized the test for its neglect in explaining how the scores should be interpreted and how the "tentative norms" were obtained. He also commented on the fact that information on reliability was based on testing only forty-nine third grade students.[92] Kingston accepted the reliability measures offered by the manual, but he stated that the tests "do not provide a measure of the reader's comprehension or interpretation of materials in context settings."[93]

---

[90] Ibid., p. 13.

[91] Ibid., p. 14.

[92] Emery P. Bliesmer, "The Silent Reading Diagnostic Tests," The Sixth Mental Measurements Yearbook, p. 1119.

[93] Albert J. Kingston, "The Silent Reading Diagnostic Tests," The Sixth Mental Measurements Yearbook, pp. 1119-20.

The authors of the Silent Reading Diagnostic Tests claimed that it gives a "complete diagnosis." They also stated that oral reading and silent reading abilities were different; yet the tests were not designed to test oral reading, nor is a measure of oral reading included on the Graphic Profile.[94] While the test authors called it a complete diagnosis, they designed it to be used only in conjunction with a silent reading test.[95] A final problem was found in considering the relevance of the grade equivalent norms for some of the subtests.

An illustration of the problem the test might pose in the classroom was discovered in considering the method used in phonic reader systems, discussed later in this paper. These readers present letter sounds the first year in school. One Silent Reading Diagnostic Tests subtest was titled Letter Sounds. The teacher was instructed to make the sounds in isolation for consonants and vowels. Thirty sounds were designed to be read and the correct number achieved by the student recorded on the Graphic Profile.[96] A score of twenty-eight correct from the possible thirty would be recorded at grade equivalent 6.0. A child could fail to

---

[94]Bond et. al., Manual, p. 2.

[95]Ibid., p. 9.

[96]Ibid., p. 7.

recognize all the vowel sounds tested and receive a grade equivalent score of 3.0.[97]

Rather than limiting itself to either oral or silent reading, the Durrell Analysis of Reading Difficulty was written with graded passages to be read orally and silently. Dallman called testing both types of reading its chief value.[98]

Like the Gray Oral Reading Tests, the Analysis was designed as an individual diagnosis. Donald D. Durrell, its author, combined subtests with standardized norms with checklists without norms. Durrell stated that the Analysis had its main purpose in remedial diagnosis. He asserted that the informal checklists were more important to this purpose than were the norms:

> The check lists for recording observations of difficulties are the most important feature of the Analysis. While norms are provided for many of the tests, a record of the difficulties the child displays is more important than the level of attainment. The items on the check lists are those of highest frequency and significance in remedial work.[99]

The Analysis was built with the following subtests:

---

[97] Ibid., p. 15.

[98] Martha Dallman, Teaching the Language Arts in the Elementary School (in Brown Education Series, ed. Lowry W. Harding. Dubuque, Iowa: Wm. C. Brown Company Publishers, 1966), p. 292.

[99] Donald D. Durrell, Manual of Directions, Durrell Analysis of Reading Difficulty (New Edition; New York: Harcourt, Brace & World, Inc., 1955), p. 3.

Oral Reading Tests, Silent Reading Tests, Listening Comprehension Tests, Supplementary Paragraphs (to be used in retesting or for additional testing), Word Recognition and Word Analysis, Visual Memory of Word Forms, Auditory Analysis of Word Elements, Spelling and Handwriting, Letter Recognition, Phonics, and Learning Rate.[100] Checklists, a Profile Chart and General History Form were also included to aid in total diagnosis, along with detailed information on how to interpret the results of the tests.[101]

One writer called the Analysis the reading test that attempted to do for reading what the Stanford-Binet did for intelligence testing but stated that it did not quite succeed.[102] Harris and Heilman pointed out that the Analysis oral reading standards, as well as silent reading standards or norms, on this particular test, were based entirely upon rate of reading, which the administrator obtains by using a stop watch. Such limited standards were found to hamper test interpretation and use.[103]

Durrell did not, however, claim that his test would give the teacher instructional level:

---

[100] Ibid., pp. 3-5.

[101] Ibid., pp. 4-5.

[102] J. Roy Newton, Reading in Your School (New York: McGraw-Hill Book Company, Inc., 1960), p. 110.

[103] Harris, How to Increase Reading Ability, p. 205; and Heilman, Principles and Practices of Teaching Reading, p. 186.

> There are many habits and skills in reading which either cannot as yet be measured satisfactorily by standard tests or which can be measured as well or better by informal tests.[104]

In addition to reading interests and effect and speed of reading in class, word skills, and speed and accuracy in finding information in textbooks, Durrell stated that the level of reading material used in the classroom could best be fitted to the child by an informal test. He described the informal test which he recommended:

> Suitability of the level of reading materials used in the classroom. Listen to the child's reading from the textbook he is using and note the number of errors on difficult words. If the child makes more than one error in twenty running words, the material is too difficult and a lower level of reading is recommended.
>
> In intermediate grades, the suitability of the textbook can be determined in a rough way by selecting the eight or ten hardest words on each of three pages and checking whether the child can read them and understand their meanings. If more than five words per page are unknown, it is likely that the material is too difficult.[105]

Durrell also noted the necessity for determining the suitability of material in the remedial as well as regular classroom situation. After administering and scoring the Analysis, the teacher was advised to use the informal method described above to suit instructional material to

---

[104] Durrell, Manual, p. 27.

[105] Ibid.

the student. The Analysis was to be used in planning how skills should be taught and faulty habits corrected.[106]

The author of the Analysis claimed to have wide experience in devising his test in particular. While he preferred the "check lists of errors" to the norms, he stated that the norms were "based on no fewer than a thousand children for each test." Furthermore, they claimed to "check satisfactorily against other measures of reading ability."[107]

Of the three types of diagnostic tests considered, the Gray Oral Reading Tests and Durrell Analysis of Reading Difficulty were designed to present the child with passages of increasing difficulty to be read orally. The author of the Analysis also included increasingly difficult passages to be read silently. The Silent Reading Diagnostic Tests included very little connected reading, and the purpose of the tests was not to establish a level of passage difficulty for suiting materials to a pupil's reading level. No proof was found that any one of the three tests would yield a realistic instructional level.

Informal reading inventories. Neither the standardized reading tests nor the diagnostic tests claimed

---

[106]Ibid., pp. 30-31.

[107]Ibid., p. 32.

that they actually give exact instructional levels. They have been criticized for this reason by many writers. Because of this failure, informal methods of determining instructional level have been used. Morton Botel summarized the testing controversy thus:

> Standardized testing methods have not been successful in determining proper placement, or "the right book." Between 10 and 15 million pupils in our schools are reading books that are too difficult for them. This estimate is based upon widespread agreement among reading specialists, who have found that one fourth to more than one third of the pupils in a typical classroom are trying to read books so difficult for them that these books actually interfere with reading progress.[108]

The originator of the informal test, the informal reading inventory, with most universally accepted standards was Emmett Betts.[109] His standards were accepted by Botel but changed because Botel used a different testing form.[110] Spache accepted the concept of the informal reading inventory but changed the percentages which he felt were too strict.[111]

Betts felt that the teacher or clinician was concerned

---

[108]Botel, loc. cit.

[109]Betts, Foundations of Reading Instruction, pp. 438-87.

[110]Botel, op. cit., pp. 5-7.

[111]George D. Spache, Reading in the Elementary School (Boston: Allyn and Bacon, Inc., 1964), pp. 242-59.

with learning about each student four stages in his reading ability. First, the teacher would want to know the highest reading level, called by Betts the basal, or independent reading level, at which a pupil can read "with full understanding and freedom from mechanical difficulties." Next, the instructional level, or level at which "systematic instruction can be initiated" should be ascertained. A third level was called the frustrational level. At this level, the reading matter is too difficult; the individual is "thwarted or baffled by the language (i.e., vocabulary, structure, sentence length) of the reading material." The fourth level Betts called the probable capacity level, or the level at which the pupil can comprehend material which he hears read. Betts intended that level be interpreted as levels of basal readers.[112]

In addition to other assumptions concerning classroom reading instruction, Betts assumed that the following criteria would establish the levels as he had defined them. In his description, briefly, the basal level was characterized by 90 per cent comprehension and 99 per cent accuracy in word recognition. There should be no signs of difficulty or tension in silent or oral reading and silent reading should render a higher

---

[112]Betts, Foundations of Reading Instruction, pp. 438-39.

comprehension score than oral reading.[113]

Betts cautioned that instructional level, even when established in accordance with his standards, should be considered only an estimated, or probable level. In spite of the possible range of error, he stated that a teacher should make the attempt to find this level, lest the pupil be frustrated in learning to read. The minimum standards at instructional level were described as a comprehension score of 75 per cent, using both factual and inferential questions; word recognition accuracy of 95 per cent in connected reading; and the other signs of comfortable, non-frustrated reading, such as good reading posture and easy, rhythmic pronunciation in oral reading after silent reading preparation. A student should be free from undesirable habits, called signs of difficulty; some signs of difficulty were described, such as fingerpointing, head movement, nail biting, and vocalization in silent reading.[114]

Betts contended that no child should be instructed at his frustrational level for at this level "the pupil is unable to comprehend printed symbols to a reasonable degree." Comprehension of less than 50 per cent and word recognition of 90 per cent or less were said to be accompanied by any number of signs of reading difficulty

---

[113] Ibid., p. 445.

[114] Ibid., p. 447.

or frustration. Betts believed that errors made at this level tend to compound themselves and to be unrecognized by the student.[115]

The probable capacity level, based upon listening and responding to material read by the teacher, was characterized by comprehension of at least 75 per cent and the ability to pronounce accurately the words in the material heard. To meet the criteria fully, the student was expected to also be able to use the words precisely in discussing the ideas presented and to be able to supply related ideas from his own experience. A check upon structural difficulty was suggested by the statement that the pupil should be able to use as complicated a structure in speaking as that used in the written passages.[116]

Any set of carefully-graded books, basal readers, science or history books, My Weekly Reader, Current Events, etc., was suggested for the testing material itself. Betts intended that the test be carried out similar to a directed reading activity, with readiness for each passage established prior to silent reading for specific purposes. Oral reading was to be done for new purposes and should follow silent reading; however, at times, oral reading at sight was

---

[115] Ibid., p. 451.

[116] Ibid., p. 453.

recommended "as one means of appraising reading performance."[117]

The steps for a full-length informal reading inventory as suggested by Betts were oral reading at sight, silent reading guided by single answer questions, and oral rereading. Oral reading at sight from a series of short selections was proposed as the first step in order to examine the pupil's behavior without preparation, to determine rapidly the point for more detailed testing, and to allow the examiner to learn when and how abruptly faulty reading habits appear in the reading of graded material.[118]

Both silent reading and the oral reading at sight should be preceeded by motivating questions and purposes for reading similar to those used in a directed reading activity, Betts contended. He suggested that the examiner could also ask the pupil to tell all that he could remember concerning a passage in lieu of asking detailed questions. Oral rereading was to be done as a response to comprehension questions also and should exhibit improvement over oral reading at sight.[119]

Betts provided a detailed discussion of the considerations which would make an informal reading

---

[117] Ibid., pp. 455-56.

[118] Ibid., p. 457.

[119] Ibid.

inventory a more valid test, including careful selection and organization of materials for testing, methods of checking comprehension accurately, and considerations in establishing rapport.[120]

Several limitations of an informal reading inventory were discussed by Betts. An informal inventory would depend to a larger extent upon the competency of the examiner than would a standardized test. If material other than the book which the student would use in the classroom were utilized, validity would be lessened according to the degree of difference between the readability levels of the two materials. Validity would also be decreased when giving inventories from different series of textbooks, again because of differences in readability, including vocabulary, structure, and content. Reliability would depend upon the degree to which criteria are interpreted alike by different examiners. Generally, reliability would be increased when length of selections increase. As with a standardized test, reliability would be affected by the degree of rapport established for each testing.[121]

Advantages which Betts listed were the low cost of

---

[120] Ibid., pp. 457-65.

[121] Ibid., pp. 471-77.

testing materials; direct and rapid administration; validity, if the examiner uses materials in which the student would be instructed; possibility of use for group and individual testing; the learner's education in how reading is appraised; the learner's awareness of progress; the appraisal of instructural, basal, frustration, and capacity levels, as well as of diagnosis of specific needs; the ability of the examiner to choose interesting materials for testing; attention directed toward the concept of graded readability; and incorporating testing with instruction without lessening the value of either.[122]

Betts created an informal reading inventory to the Betts Basic Readers, described later in this paper. Albert J. Harris, Nila Banton Smith, and Robinson and Rauch were found among the writers using the Betts standards for informal evaluation of reading levels.[123]

Most of the comparisons between reading tests used the informal reading inventory, with Betts standards, as the criterion test. Robert Karlin declared that true

---

[122] Ibid., pp. 478-79.

[123] Harris, Effective Teaching of Reading, pp. 120-25; Smith, Graded Selections for Informal Reading Diagnosis, xiii; and H. Alan Robinson and Sidney J. Rauch, Guiding the Reading Program (Chicago: Science Research Associates, 1965), pp. 72-74.

instructional levels should be established for children. He stated:

> If we are to reduce or eliminate the reading difficulties elementary school pupils experience, we must determine what they are. In addition, we must know at what levels to offer instruction and the levels at which pupils can read books without help from us. Most schools administer standardized reading tests for children in order to find out at what level they are reading and what their reading needs and strengths are.[124]

To ascertain the relevance of such testing programs, Karlin surveyed research on tests. In studies by Williams, McCracken, Sipay, Kress, and Long, the informal inventory was used as criterion by the writers, and the standardized tests were said to overestimate instructional level. Karlin found no studies and only one book in which the authors preferred the standardized tests over informal inventories.[125] The book was Reading Difficulties: Their Diagnosis and Remediation by Bond and Tinker who had claimed:

> The diagnosis of reading ability and deficiencies is best achieved through using standardized tests and

---

[124]Robert Karlin, "Remedial Studies at Elementary Level: Their Implications for Classroom Practice," Combining Research Results and Good Practice, Mildred A. Dawson, editor. Proceedings of the Eleventh Annual Convention of the International Reading Association (Newark, Delaware: International Reading Association, 1967), p. 102.

[125]Ibid., pp. 101-8.

> procedures. It is possible, however, and sometimes desirable to make a preliminary diagnosis by less formal procedures.[126]

They then described the informal inventory, using Betts' standards. They commented that important information is gained thus, but not enough diagnostic detail unless other procedures are included.[127] They concluded that, although the informal diagnosis "explores some areas not treated in standardized procedures,"

> Obviously the more systematic, standardized diagnostic procedures . . . are more reliable and yield a more thorough diagnosis than any informal procedure, especially with the more severe disability cases.[128]

Bond and Tinker did not recommend any one test which gave accurate instructional level, with the exception of the informal inventory, in spite of their preference for the standardized tests for total diagnosis. They suggested that the average reading grade from several tests could be used but placed emphasis entirely upon diagnosis of skill development.[129]

From the investigations of studies related to

---

[126]Bond and Tinker, op. cit., p. 103.

[127]Ibid., pp. 198-200.

[128]Ibid., p. 203.

[129]Ibid., pp. 195-234.

instructional level, Karlin concluded:

> Might I suggest that we have some evidence to support the view that suitable informal reading tests will yield results which will enable use to more adequately provide reading instruction for those who need it badly. This recommendation does not preclude the use of standardized reading tests for evaluating the reading of poor readers. It merely underscores other possibilities available to us as we work with children who require special attention.[130]

A similar conciliatory view was found in a study by Putt and Ray. They outlined the failures of which both standardized tests and informal tests have been accused: standardized tests, of overestimating performance; and informal tests, of lacking reliability, validity, and unified skill measurement. Putt and Ray examined both kinds of tests in the classroom to determine their actual roles.[131]

Twenty-five fifth grade students at the Campus School of State University College, Fredonia, New York, were the subjects. The intelligence quotients for these middle- or upper-middle class children ranged from 97 to 135, and 60 per cent of the pupils had one or more parent with college background. Their reading achievement had been carefully recorded prior to the study.

---

[130]Karlin, op. cit., p. 102.

[131]Robert C. Putt and Darrel D. Ray, "Putting Test Results to Work," The Elementary School Journal, LXV (May, 1965), 439-444.

Three standardized tests were used, including the Wide Range Achievement Test (Reading Section); the Gates Reading Survey, Form I; and Silent Reading Diagnostic Tests, Form D-A. The third test was administered to only those students whose grade equivalent score on the Gates test was below 5.0. The Betts standards were used to construct an informal reading inventory from original material "similar to that in the basal reader used in the classroom."[132] Readability of the passages were checked by using the Spache and Dale-Chall formulae.

Results from the two types of tests were superimposed on a graphic profile upon which the child's reading expectancy level (using the Bond formula) and the class average were drawn. The authors stated that the profiles enabled the teacher to group her class more successfully and to plan the skill development in areas in which the entire class showed weaknesses. She was also able to isolate individual weaknesses for flexible grouping. The authors claimed that each type of measurement was needed for thorough diagnosis and "adjustment of classroom instruction."[133]

---

[132] Ibid., p. 440.

[133] Ibid., p. 444.

The authors noted that, while 66 per cent of the students were reading below their potential, none were so far below that they were classified as disabled readers. Four groups resulted from the testing, at grade levels 4^2, 5^2, 6^1, and 7^1.[134] Putt and Ray advised the classroom teacher to use both types of test in compiling a diagnosis suitable for this type of classroom instruction:

> Instructional levels could not be determined from the standardized tests, but each child maintained the same relative class rank on both the informal and the standardized tests.
>
> The informal reading inventory did not reveal relative group weaknesses. Vocabulary weaknesses, for example, were determined from the standardized tests.[135]

The originator of the informal reading inventory and Carolyn M. Welch created an informal inventory to use with the Betts Basic Readers by American Book Company. The test consisted of a word recognition test and paragraphs from the readers, third preprimer through sixth reader. Comprehension questions were written to accompany the passages. Betts and Welch changed the standards somewhat for instructional level. More than 95 per cent accuracy in oral reading and freedom from "all signs of difficulty" were required as in Betts' original description of

---

[134] Ibid., p. 441.

[135] Ibid., p. 444.

instructional level, but the comprehension standard was raised to "about 80%."[136]

Of the tests described in this paper, only the proficiency tests and informal inventories claimed as a main purpose to obtain an instructional level for the student. Betts and Welch emphasize that:

> THE SPECIFIC PURPOSE of an informal reading inventory is to identify the highest reader level at which each pupil can read, under teacher supervision, without signs of difficulty. This reader level is called the instructional, or teaching, level.[137]

The test manual described the subtests as an informal word recognition inventory and an informal reading inventory. The word recognition test was attributed with two purposes, that of determining a probable starting point for the reading inventory and of measuring word perception achievement.[138] Three categories of recognition were established: flash recognition; untimed recognition of words unable to be recognized within a three second flash exposure; and phonic recognition, with the teacher asking the pupil to point to the unknown element which is then supplied. The test

---

[136]Emmett A. Betts and Carolyn M. Welch, Manual for Teachers, Informal Inventories (New York: American Book Company, 1964), p. 1.

[137]Ibid.

[138]Ibid., p. 2.

standard was devised by using only the flash score; the two remaining scores were to be used for diagnostic purposes. Two errors at any reader level at flash recognition was termed the stopping point for testing.[139]

The informal reading inventory section of the Betts inventory was recommended for yielding two levels, the instructional and independent reading levels. The administrator was not advised to continue the test above instructional level.[140] Ruth Strang commended this practice of stopping at instructional level as an important advantage of informal inventories as opposed to standardized tests, which expect a child to reach frustration level so that the grade norms will fit properly. Strang said this practice was particularly good for the poor reader accustomed to failure in school.[141]

Betts and Welch designed their inventory without passages for oral reading at sight. The procedure suggested was to prepare the student for silent reading by asking a few motivating questions or initiating any motivating conversation relevant to the passages, having the student

---

[139] Ibid., pp. 2-3.

[140] Ibid., pp. 3-4.

[141] Ruth Strang, "Teaching Reading to the Culturally Disadvantaged in Secondary Schools," Journal of Reading, X (May, 1967), 527.

read silently to answer specific purposes listed below the reading passage in the teacher's edition of the test, and requesting oral rereading for specific purposes listed beneath the Comprehension questions. No indication was given that the pupil read the selection without the interruption or guidance of the comprehension questions, nor that he could answer the questions after such a reading.[142] The Betts Basic Reader inventory appeared to depart from the procedures described in Foundations of Reading Instruction as described earlier.[143]

Another departure from the original Betts procedure was devised by Morton Botel and titled the Botel Reading Inventory. This test was divided into three sections, Word Recognition Tests, Word Opposites Tests, and Phonics Mastery Tests, for two equivalent forms. The Word Recognition Test and Word Opposites Test were called Reading Placement Tests by Botel and standards for their interpretation were provided.[144]

---

[142]Betts and Welch, Manual for Teachers, Informal Inventories, pp. 3-4.

[143]Betts, Foundations of Reading Instruction, pp. 438-87.

[144]Botel, op. cit., pp. 22-23.

Botel advised that two reading levels should be gained from the Reading Placement Tests, the instructional level and the frustration level. The two levels were described by Botel according to Betts' standards, but Botel contended that his tests, which were made up of isolated words rather than connected reading matter, required a revision of the standards. The criteria for instructional level in the Botel Reading Inventory were defined as 70-90 per cent recognition and sight pronunciation of the words on the Word Recognition Test and comprehension of 70-80 per cent of the words of the Word Opposites Test. Frustrational level was interpreted as any greater percentage of error than that acceptable at instructional level. Botel cautioned that, by changing the procedure and standards, the results would yield only estimates of reading levels.[145]

For the classroom teacher, Botel described standards which he recommended for determining "free reading," instructional, and frustration levels. His interpretation departed somewhat from the Betts' standards. He said that instructional level should not be construed as a single level, but a range of levels at which a child can read

---

[145] Ibid., pp. 6-7.

[146] Ibid., p. 22.

profitably.[147] David Russell upheld this variable-level theory, opposing using a set percentage of accuracy to judge levels.[148]

Botel has used the Betts standards but changed the percentages; he also indicated that there were times when a student could read profitably from a book above his instructional level. Betts had also offered this suggestion.[149] Botel did not advise using a basal reader at frustration level, but indicated that library books might be at this level.[150]

The Botel Reading Inventory was applauded by Ira E. Aaron for having carefully graded word lists.[151] He criticized Botel's assumption of a relationship between comprehension in connected reading and the ability to choose words opposite in meaning.[152] A more severe criticism was that concerning Botel's standards. Botel had claimed that research backed the Betts standards.[153] Aaron assumed that

---

[147] Ibid., p. 22.

[148] David H. Russell, Children Learn to Read (New York: Ginn and Company, 1949), p. 151.

[149] Betts, Foundations of Reading Instruction, p. 443.

[150] Botel, loc. cit.

[151] Ira E. Aaron, "Botel Reading Inventory," The Sixth Mental Measurements Yearbook, p. 1121.

[152] Ibid., pp. 1121-22.

[153] Botel, op. cit., p. 6.

Botel had claimed such support for his adaptation of the criteria and stated that he could find "very little research evidence" for these standards.[154]

Brown criticized the criteria, as well as the lack of normative data, reliability, validity, or even tryout data. He concluded that the placements yielded were inexact:

> . . . one wonders how this "instrument" would be any better than a simple application of informal reading inventory criteria to the oral and silent reading of graded material. The selection of appropriate instructional and free reading material on the basis of the scores on this inventory would be rather gross at best and, further, would be very dependent upon some evaluation of the readability level of those instructional materials.[155]

II. BASES UPON WHICH TESTS WERE CONSTRUCTED

Proficiency standards. Basal reader proficiency tests were based upon mastery of the vocabulary, skills, and content of the series for which the tests were designed. The Reading-Study Achievement Tests required proficiency of 95 to 97 per cent on all subtests to accompany a reader level, with no signs of difficulty in silent or oral reading.[156]

---

[154]Aaron, op. cit., p. 1122.

[155]Charles M. Brown, "Botel Reading Inventory," The Sixth Mental Measurements Yearbook, p. 1122.

[156]Betts and Welch, Overview, p. 2.

The Basic Reading Test to the Scott, Foresman readers also required that a pupil should demonstrate mastery, or a high proficiency, on the test items.[157]

The informal reading inventory as described by Betts emphasized proficiency in reading at the independent level, with 90 per cent comprehension, 99 per cent word recognition, and no signs of difficulty. The instructional level was a level of slightly lowered proficiency: no lower than 95 per cent word recognition with 75 per cent comprehension, still with no signs of difficulty. Any performance less proficient was called frustrating.[158]

George Spache disputed the Betts standards as being "arbitrarily high" and unrealistic as compared with standards achieved by an actual standardization using "widespread pupil testing and comparison with other individual and group reading tests." He did not, however, offer comparative research to support his contention.[159]

Two studies were discovered which investigated the validity of the Betts standards to a limited extent. Sipay compared standardized test results with those of two

---

[157]Robinson et. al., loc. cit.

[158]Betts, Foundations of Reading Instruction, pp. 438-39.

[159]Spache, op. cit., p. 247.

informal reading inventories using different levels of word recognition as criteria, 96 and 90 per cent, with the comprehension standard held constant. Standardized tests continued to overestimate the levels attained by both informal tests to a significant level. In his conclusions, Sipay recommended that teachers should know which test and criteria are being used when instructional levels are given for children. He did not judge between the two informal standards.[160]

Helen Robinson reported a study by Elder on word recognition standards. Elder claimed that students could tolerate as few as 88 per cent known words when he substituted nonsense words for key words; however, his results were based upon student answers to yes/no questions concerning how difficult they found the material. He did not ask comprehension questions to learn how much of the material was understood.[161]

Norming populations. Rather than employing static proficiency standards, standardized reading survey and diagnostic tests were based upon norming populations.

---

[160]Edward R. Sipay, "A Comparison of Standardized Reading Scores and Functional Reading Levels," The Reading Teacher, XVII (January, 1964), 265-68.

[161]Helen M. Robinson, "Summary of Investigations Relating to Reading, July 1, 1962, to June 30, 1963," The Reading Teacher, XVII (February, 1964), 326.

They intended to present the reading situation that actually occurs among the normal range of children. Groups of children who supposedly represent the average range of abilities have been administered tests and norms established from the results, usually after statistical manipulations to assure that a normal distribution of scores were obtained. These norms have been published along with test manuals, so that a teacher might compare the performance of ner students with that of others of the same grade, age, and, hopefully, socio-economic background; thus test scores have been given in terms of grade equivalents describing the norming population.[162]

Strang explained the background for criticism of the grade equivalents in overestimating instructional level. She stated:

> These scores show the level at which comprehension can be achieved with difficulty rather than the level at which fluency and reasonable accuracy can be expected.[163]

The difficulty of grouping by using such norms as standards was pointed out by several writers. An

---

[162]Lee J. Cronbach, Educational Psychology (New York: Harcourt, Brace & World, Inc., 1954), p. 185.

[163]Ruth Strang, Reading Diagnosis and Remediation. An IRA Research Fund Monograph (Newark, Delaware: International Reading Association Research Fund, 1968), p. 95.

illustration was furnished by the Balow and Curtin study in which homogeneous grouping was attempted by using the comprehension subtest scores from a standardized test. The resulting groups were heterogeneous in other skills, the authors asserted, even though the subtest used indicated that they would be homogeneous.[164]

Roma Gans asserted that teachers tend to become embarrassed or affronted when students fall below grade level, even though approximately half of the class should fall above and half below the grade level. Gans declared that this false expectation forms the basis for much money wasted on classrooms of textbooks all "on grade level."[165]

Analyses of reading skills. Diagnostic reading tests had sections based upon analyses of reading skills at different developmental reading levels. Basal reader proficiency tests also advised the use of checklists of skills and had subtests on skills presented at different basal reader levels.

The Silent Reading Diagnostic Tests presented a subtest, discussed earlier in this paper, called Letter

---

[164]Bruce Balow and James Curtin, "Reading Comprehension Scores as a Means of Establishing Homogeneous Classes," The Reading Teacher, XIX (December, 1965), 169-73.

[165]Roma Gans, "Misspent Funds and the Consequences," The Reading Teacher, XX (April, 1967), 595-99.

Sounds which was accused of having an unrealistic grade equivalent scale.[166] Bond and Tinker presented a detailed summary of the different skills which are tested by various diagnostic tests, illustrating the variety of skills and the differing emphasis given the skills by test-writers.[167] The Durrell Analysis was cited for exempting its checklists of skill difficulties from the norming procedure.[168]

Because of difficulty in establishing tests for the skills that generalize to other reader series, basal reader proficiency tests were considered to have limited utility. The Reading-Study Achievement Tests authors pointed out the possibility, even, that skills might be taught differently or omitted from their basal readers by a local school system.[169]

Analyses of reading skills have been found to vary according to the philosophy upon which the reading program has been adopted. For example, Jeanne Chall found problems in deciding which reading system was better for students because tests would favor one system or another due to variation in when skills are presented

---

[166]Bond et. al., Teacher's Manual to Silent Reading Diagnostic Tests, p. 15.

[167]Bond and Tinker, op. cit., pp. 203-37.

[168]Durrell, op. cit., p. 3.

in different reading systems.[169] She suggested "one test for one skill" as a replacement.[170] Other writers have suggested checklists for reading levels, independent of any one reading program.[171]

III. RELATION OF STANDARDS TO METHODS OF TEACHING READING

"Whole-word" basal readers. Jeanne Chall laid to rest the contention that today's basal readers can be divided into "look-say" readers that teach no phonics and phonic readers that do. Rather, basal readers since 1930 were found to fall into two main categories, those that teach phonics "intrinsically" after a body of sight words has been taught, and those that present phonics systematically from the beginning of instruction.[172] This false dichotomy was discussed by Heilman in his book advising a system which

---

[169]Betts and Welch, Overview, pp. 1-2.

[170]Jeanne Chall, Learning to Read: The Great Debate (in The Carnegie Series in American Education. New York: McGraw-Hill Book Company, 1967), pp. 103-8.

[171]Ibid., pp. 212-13.

[172]DeBoer and Dallman, op. cit., pp. 320-24; and Walter B. Barbe, Educator's Guide to Personalized Reading Instruction (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1961), pp. 142-42; 152-53, etc. Barbe does include sight words which might not be fair to a phonic reader system.

he contended was a sensible phonic basis for any series of readers.[173]

Most writers contended that the standardized survey and diagnostic tests could not yield a score that reflected true instructional level for these basal readers. Paul McKee, however, expressed a belief that standardized test scores might be used for tentative grouping.[174] Jeanne Chall suggested that the standardized test technique might be more universally trusted if "single-component tests" were used rather than survey tests so that mastery of individual skills could be assessed.[175]

In order that tests would tell teachers more about a child's ability to read a book, Whipple and other writers investigating difficulty level in basal readers encouraged educators to urge textbook publishers to answer a standard questionnaire concerning their books, in an attempt to apply a uniform measure to all basals. She did not furnish such a questionnaire, however.[176]

---

[173]Arthur W. Heilman, Phonics in Proper Perspective (Columbus, Ohio: Charles E. Merrill Books, Inc., 1964), p. 2.

[174]Paul McKee, Reading/A Program of Instruction for the Elementary School (Boston: Houghton Mifflin Company, 1966), p. 250.

[175]Chall, loc. cit.

[176]Gertrude Whipple, Upgrading Elementary Reading Programs. Citation Address of the Eleventh Annual Convention of the International Reading Association (Newark, Delaware: International Reading Association, 1966), p. 4.

"Whole-word" editors have written proficiency tests and manuals which stress a mastery of the skills and content in the readers. These tests and the informal reading inventory were found to be recommended as appropriate means of establishing a child's instructional level in "whole-word" series.

Phonic readers. Phonic readers were designed as a result of the definition of reading as decoding, or "turning the printed symbols into the sounds which are language," as a basic skill "to get at the meaning."[177] Chall contended that silent reading tests penalized children instructed by this method when the test was administered in the first grade, because standardized silent reading tests were timed and written with many of the irregular "sight words."[178]

Authors of phonic readers admittedly do not put emphasis upon fluency and mastery at the early stages of instruction, thus even the informal reading inventory would penalize the reader more than the authors intended. They stated that the teacher should not wait for fluency or mastery but should "move along rather rapidly through the

---

[177]Glenn McCracken and Charles C. Walcutt, Basic Reading, Teacher's Edition for Pre-Primer and Primer (New York: J. B. Lippincott Company, 1963), iv.

[178]Chall, op. cit., pp. 106-7.

Grade One program, so that her class will complete it in one school year."[179] No proficiency tests were written to accompany the Lippincott series, but the Grade One skills were designed to be reviewed at the beginning of the second grade, "so that children may move rapidly into free reading in their second school year."[180]

The criticisms offered against the survey and diagnostic tests applied to phonic readers as well as "whole-word" readers. Informal reading inventory standards were designed to fit any connected reading material, including that in phonic readers. Many writers asserted that readability standards should be applied to the passages selected to assure that the increase in difficulty is regular and even.[181] If such a readability check were conducted, informal inventories would fit phonic readers, with the exception of the lower grades as indicated above. The informal inventory would only fit the early stages of reading instruction in a series like Lippincott

---

[179]McCracken and Walcutt, op. cit., p. viii.

[180]Glenn McCracken and Charles C. Walcutt, "Basic Reading, Grades 1 through 8," advertising pamphlet (Philadelphia: J. B. Lippincott Company, undated), p. 2.

[181]George H. McGinnis, "The Readability Graph and Informal Reading Inventories," The Reading Teacher, XXII (March, 1969), 516-522.

if the Lippincott philosophy were changed.

Individualized reading. Three terms used as a catch-phrase for individualized reading, "seeking, self-selection, and pacing," were discussed in relation to the concept of instructional level by some writers, for these words imply that a child can be taught to choose his own reading material for instruction. Furthermore, writers like Veatch asserted that a child taught thus can move to more difficult material through self-selection without violating developmental reading principles.[182]

Standardized survey and diagnostic tests did not claim to help a teacher provide library books on the proper instructional level for children. In fact, with the individualized reading system in which children choose their own books, the teacher was instructed not to "place a child in a book." Instead, Jeannette Veatch, the leading proponent, asserted that a child should be taught to test himself in the book which he finds interesting. She said:

> At the outset, children may need help in choosing their reading material wisely. They should be encouraged to choose a book they can read and one that they will like. If more explanation is needed, it should be given. Finding out if it is a book one is able to read means skimming a page or two to see whether there are too many hard words. We can use as a guide "three

---

[182]Jeannette Veatch, Individualizing Your Reading Program (in The Putnam Series in Education, eds. Ole Sand and Blaine Cook. New York: G. P. Putnam's Sons, 1959), pp. 4-7.

unknown words to a page," but this is an approximate number.[183]

This testing method employed word recognition standards similar to the informal reading inventory and used Betts' technique of informing the student of the standards used.[184] Comprehension standards were ignored, for the originators of individualized reading asserted that comprehension standards could be strained under the pull of motivation, or interest.[185] Thus Veatch reiterated the criteria she stated earlier: "a book they can read and one that they will like."[186]

Many criticisms have been offered against the use of "interest" as a comprehension standard. Carillo found that poorer readers would not select books realistically.[187] Carlton and Moore contended that not interests, but self-concept rules the books that a child selects.[188] Hicks and Simpson,

---

[183] Ibid., pp. 19-20.

[184] Betts, Foundations of Reading Instruction, p. 462.

[185] Veatch, op. cit., pp. 20-21.

[186] Ibid., p. 19.

[187] Lawrence Carillo, "The Reading Program, K-6," SRA Reading Institute Extension Service, Unit One (Chicago: Science Research Associates, Inc., 1967), p. 23.

[188] Lessie Carlton and Robert H. Moore, Reading, Self-Directive Dramatization and Self-Concept (in International Series in Education, ed. Kimball Wiles. Columbus, Ohio: Charles E. Merrill Publishing Company, 1963), p. 7.

in a study with Ira Aaron and Byron Callaway, concluded that the role of interests had either been overrated or poorly tested.[189]

If basal readers were used with individualized reading, the problem of instructional level for the basal reader would be able to be solved by the methods discussed earlier. In studying how much individualized reading is being done in this country today, Dayton G. Rothrock found that, while most teachers thought individualized reading very important, few used it alone as a method. Rothrock found that most teachers fused individualized principles with others into "an eclectic plan," many combining it with basal reader instruction.[190] Dechant said that the trend in reading is toward such an eclectic system, with no one method idolized.[191]

Thus, the informal reading inventory was found to be relevant to placing a child properly in an individualized, or partly individualized system. Veatch used similar

---

[189]Ira E. Aaron, Byron Callaway, Randall Hicks, and Hazel D. Simpson, "Reading Achievement in a Summer Reading Program," Elementary English, XLIV (December, 1967), 875-77.

[190]Dayton G. Rothrock, "Teachers Surveyed: A Decade of Individualized Reading," Elementary English, XLV (October, 1968), 757.

[191]Emerald V. Dechant, Improving the Teaching of Reading (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1964), p. 235.

word recognition standards, but no comprehension standards. Thus the comprehension standards would be superfluous, according to Veatch, due to the motivation of interests.

CHAPTER III

SUMMARY AND CONCLUSIONS

In this paper, selected systems of teaching reading were examined to determine what testing procedures their originators recommended for placing a child in proper instructional level. The testing tools and their foundations were also discussed. Certain conclusions were obtained with regard to ascertaining a child's instructional level.

I. SUMMARY

Instructional level was defined as the level of book which the child can read with success in word recognition and comprehension, yet which contains a degree of unlearned material (vocabulary, concepts, or skills). Four types of reading tests were selected for examination regarding their competence in aiding teachers to determine a child's instructional level.

Basal reader proficiency tests were designed to show a student's progress in learning the vocabulary, skills, and content of a particular series of graded readers. Such tests were based upon the belief that the student should exhibit near-perfect performance

on a test accompanying a book before being instructed in the next book in the series.

The Reading-Study Achievement Tests, written to parallel the Betts Basic Readers by the American Book Company, were examined as a sample of basal reader proficiency test. The editors required scores of 95 to 97 per cent accuracy on all subtests, with no signs of difficulty in silent reading, before a student was to be classified as ready for the next book. Such tests were called standard tests by the authors, as opposed to standardized tests.

Standardized reading achievement tests, such as the Gates-MacGinitie Reading Tests, were prepared to inform a teacher how well a student reads certain passages in comparison with age and grade peers. Scores from standardized tests generally were presented in terms of grade equivalents, indicating the reading level at which a student reaches complete failure. Such tests were called survey tests. They normally had lengthy subtests (usually only two, vocabulary and comprehension) which tended to increase reliability on the aspects tested.

While some writers claimed that standardized tests could yield instructional levels for students, most writers contended that such tests overestimate instructional levels extremely, particularly instructional levels of poor readers who tend to mark answers at random rather than be humiliated by failure early in the testing session.

Studies illustrated that scores from such random markings varied widely according to the grade level of the test administered.

Three types of diagnostic reading tests were examined, an oral reading test, a silent reading test, and a test which measured both types of reading. Of these three types of test, only the example of the silent reading diagnostic test claimed to yield a complete diagnosis. The Gray Oral Reading Tests were held to penalize unduely the readers at beginning grade levels. The Silent Reading Diagnostic Tests did not test oral reading and depended upon a number of other measures to yield the complete diagnosis which it purported to provide. The author of the Durrell Analysis of Reading Difficulty, a test of both oral and silent reading, admitted that his test could not predict what level of reading material would best suit the abilities of a child. He recommended the use of an informal test similar to the informal reading inventory.

The fourth type of test to determine instructional level was the informal reading inventory. Most writers contended that the only valid means whereby instructional level could be obtained was by having a child read a series of carefully graded passages until he hit a level at which there was a reasonable amount of material unknown to him.

The most widely-accepted standards for the informal reading inventory were those extended by Emmett Betts in 1946. Betts stated that four levels were of concern to the teacher: the independent level, the instructional level, the frustrational level, and the probable capacity level. The standards for instructional level were word recognition accuracy of at least 95 per cent, comprehension of at least 75 per cent, and no signs of difficulty in silent or oral reading. The testing procedure was established to parallel the normal procedure of classroom reading instruction, the directed reading activity.

Reading tests were found to be based upon three types of standards: proficiency standards, normative populations, and analyses of reading skills. Only the proficiency standard was presented as relevant to the idea of instructional level determination. Two types of test were based upon proficiency standards: the basal reader proficiency test and the informal reading inventory.

Three methods of teaching reading were examined. Only basal reader proficiency tests and informal reading inventories were found relevant to instructional level. The basal reader proficiency test was limited to the basal reader series for which it was written. The informal reading inventory was relevant to all three types. The

inventory as described by Betts determined instructional level in the "whole-word" or intrinsic phonics readers with no difficulty; it presented some problems with phonic readers because such books provided little connected reading matter in early instruction and did not expect fluency early. The informal inventory standards were partially adopted by the popularizers of the individualized reading program. The word recognition standards were recommended; however, the comprehension standards were largely ignored. The individualized reading method contended that comprehension would be a by-product of the child's interest in materials chosen by himself and discussed in private conference with his teacher.

II. CONCLUSIONS

The informal reading inventory was found to be the only generally accepted test that would accurately yield a child's instructional level. This type of test was used as criterion test in studies examining standardized survey or diagnostic tests. The Betts standards were almost universally accepted, and no convincing disproof of his standards was offered.

Although experts attested to the utility of the Betts standards, studies were lacking (but highly recommended) that investigated whether placement of a child in a reader

at which he has 95 per cent success in word recognition and 75 per cent success in comprehension actually leads to better progress than placement in a reader using other percentages of proficiency for the informal reading inventory.

BIBLIOGRAPHY

A. BOOKS

Barbe, Walter B. Educator's Guide to Personalized Reading Instruction. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1961.

Betts, Emmett Albert. Foundations of Reading Instruction. New York: American Book Company, 1946.

_____ and Carolyn M. Walch. Betts Basic Readers, Teacher's Guide for the Fourth Reader. Third edition, revised. New York: American Book Company, 1963.

_____ _____ _____. Manual for Teachers, First-Reader Level, Reading-Study Achievement Tests. New York: American Book Company, 1963.

_____ _____ _____. Manual for Teachers, Informal Inventories. New York: American Book Company, 1964.

_____ _____ _____. Overview for Teachers, Reading-Study Achievement Tests. New York: American Book Company, 1964.

Bond, Guy L., Theodore Clymer, and Cyril J. Hoyt. Teacher's Manual to Silent Reading Diagnostic Tests. Chicago: Lyons & Carnahan, 1955.

_____ and Miles A. Tinker. Reading Difficulties: Their Diagnosis and Correction. Second edition. New York: Appleton-Century-Crofts, Inc., 1967.

Botel, Morton. Guide to the Botel Reading Inventory. Chicago: The Follett Publishing Company, 1966.

Buros, Oscar Krisen (ed.). The Sixth Mental Measurements Yearbook. Highland Park, New Jersey: The Gryphon Press, 1965.

Cronbach, Lee J. Educational Psychology. New York: Harcourt, Brace & World, Inc., 1954.

DeBoer, John J., and Martha Dallman. The Teaching of Reading. Revised edition. New York: Holt, Rinehart and Winston, Inc., 1965.

Dechant, Emerald V. Diagnosis and Remediation of Reading Disability. West Nyack, New York: Parker Publishing Company, Inc., 1968.

_____. Improving the Teaching of Reading. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1964.

Durrell, Donald D. Manual of Directions, Durrell Analysis of Reading Difficulty. New edition. New York: Harcourt, Brace & World, Inc., 1955.

Gates, Arthur I. Manual for the Gates Advanced Primary Reading Tests. New York: Teachers College Press, 1961.

_____ and Walter H. MacGinitie. Teacher's Manual, Gates-MacGinitie Reading Tests. New York: Teachers College Press, 1965.

_____ _____. Technical Manual, Gates-MacGinitie Reading Tests. New York: Teachers College Press, 1965.

Gray, William S. Examiner's Record Booklet for the Gray Oral Reading Tests. New York: The Bobbs-Merrill Company, Inc., 1963.

_____. Manual to Gray Oral Reading Tests, Helen M. Robinson, editor. New York: The Bobbs-Merrill Company, Inc., 1963.

Harris, Albert J. Effective Teaching of Reading. New York: David McKay Company, Inc., 1962.

_____. How to Increase Reading Ability. Fourth edition, revised. New York: David McKay Company, Inc., 1961

Heilman, Arthur W. Phonics in Proper Perspective. Columbus, Ohio: Charles E. Merrill Books, Inc., 1964.

_____. Principles and Practices of Teaching Reading. Columbus, Ohio: Charles E. Merrill Publishing Company, 1961.

Lorec, M. Ray. Psychology of Education. New York: The Ronald Press Company, 1965.

McCracken, Glenn, and Charles C. Walcutt. Basic Reading, Teacher's Edition for Pre-Primer and Primer. New York: J. B. Lippincott Company, 1963.

McKee, Paul. Reading/A Program for Instruction for the Elementary School. Boston: Houghton-Mifflin Company, 1966.

Newton, J. Roy. Reading in Your School. New York: McGraw-Hill Book Company, Inc., 1960.

Robinson, H. Alan, and Sidney J. Rauch. Guiding the Reading Program. Chicago: Science Research Associates, Inc., 1965.

Robinson, Helen M., Marion Monroe, and A. Sterl Artley. Guidebook, More Friends Old and New. New York: Scott, Foresman and Company, 1963.

Russell, David H. Children Learn to Read. New York: Ginn and Company, 1949.

Smith, Nila Banton. American Reading Instruction. Second edition. Newark, Delaware: International Reading Association, 1965.

_____. Graded Selections for Informal Reading Diagnosis, Grades 4 through 6. New York: New York University Press, 1963.

_____. Reading Instruction for Today's Children. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963.

Spache, George D. Reading in the Elementary School. Boston: Allyn and Bacon, Inc., 1964.

Yoakam, Gerald A. Basal Reading Instruction. New York: McGraw-Hill Book Company, Inc., 1955.

B. BOOKS: PARTS OF SERIES

Carlton, Lessie, and Robert H. Moore. Reading, Self-Directive Dramatization and Self-Concept. In International Series in Education, ed. Kimball Wiles. Columbus, Ohio: Charles E. Merrill Publishing Company, 1968.

Chall, Jeanne. Learning to Read: The Great Debate. In The Carnegie Series in American Education. New York: McGraw-Hill Book Company, 1967..

Dallman, Martha. Teaching the Language Arts in the Elementary School. In Brown Education Series, ed. Lowry W. Harding. Dubuque, Iowa: Wm. C. Brown Company Publishers, 1966.

Veatch, Jeannette. Individualizing Your Reading Program. In The Putnam Series in Education, eds. Ole Sand and Elaine Cook. New York: G. P. Putnam's Sons, 1959.

Wilson, Robert M. Diagnostic and Remedial Reading. In Merrill's International Education Series, ed. Kimball Wiles. Columbus, Ohio: Charles E. Merrill Publishing Company, 1967.

C. PUBLICATIONS OF THE GOVERNMENT, LEARNED SOCIETIES, AND OTHER ORGANIZATIONS

Karlin. Robert. "Remedial Studies at Elementary Level: Their Implications for Classroom Practice," Combining Research Results and Good Practice, Mildred A. Dawson, editor. Proceedings of the Eleventh Annual Convention of the International Reading Association. Newark, Delaware: International Reading Association, 1967.

Strang, Ruth. Reading Diagnosis and Remediation. An IRA Research Fund Monograph. Newark, Delaware: International Reading Association Research Fund, 1968.

Whipple, Gertrude. Upgrading Elementary Reading Programs. Citation Address of the Eleventh Annual Convention of the International Reading Association. Newark, Delaware: International Reading Association, 1966.

D. PERIODICALS

Aaron, Ira E., Byron Callaway, Randall Hicks, and Hazel D. Simpson. "Reading Achievement in a Summer Reading Program," Elementary English, XLIV (December, 1967), 875-77, 882.

Balow, Bruce, and James Curtin. "Reading Comprehension Score as a Means of Establishing Homogeneous Classes," The Reading Teacher, XIX (December, 1965), 169-73.

Chall, Jeanne, and Shirley Feldmann. "First Grade Reading: An Analysis of the Interactions of Professed Methods, Teacher Implementation and Child Background," The Reading Teacher, XIX (May, 1966), 569-75.

Clare, George R. "The Role of Word Frequency in Readability," Elementary English, XLV (January, 1968), 12-22.

Cleland, Donald L. "Clinical Materials for Appraising Disabilities in Reading," The Reading Teacher, XVII (March, 1964), 428-34.

Cohn, Stella M., and Margaretta W. Fite. "Personal-Social Changes Reflected in Reading Accuracy Ratings," The Reading Teacher, XVII (November, 1963), 97-99.

Coleman, E. B. "Experimental Studies of Readability," Elementary English, XLV (February, 1968), 166-78.

Dykstra, Robert. "The Effectiveness of Code- and Meaning-Emphasis Beginning Reading Programs," The Reading Teacher, XXII (October, 1968), 17-23.

Edwall, Eldon E., and Ida Bell Henry. "How to Find Books Children Can Read," The Reading Teacher, XXII (December, 1968), 230-32.

Gans, Roma. "Misspent Funds and the Consequences," The Reading Teacher, XX (April, 1967), 595-99.

Hardin, Veralee B., and Wilbur S. Ames. "A Comparison of the Results of Two Oral Tests," The Reading Teacher, XXII (January, 1969), 329-34.

Hunter, Madeline. "The Dimensions of Nongrading," The Elementary School Journal, LXV (October, 1964), 20-25.

McGinnis, George H. "The Readability Graph and Informal Reading Inventories," The Reading Teacher, XXII (March, 1969), 516-22.

Putt, Robert C., and Darrel D. Ray. "Putting Test Results to Work," The Elementary School Journal, LXV (May, 1969), 439-44.

Robinson, Helen M. "Summary of Investigations Related to Reading, July 1, 1962, to June 30, 1963," The Reading Teacher, XVII (February, 1964), 326-92.

Rothrock, Dayton G. "Teachers Surveyed: A Decade of Individualized Reading," Elementary English, XLV (October, 1963), 754-57.

Sava, Samuel G. "When Learning Comes Easy," Saturday Review, LI (November 16, 1968), 102-4, 119.

Sipay, Edward R. "A Comparison of Standardized Reading Scores and Functional Reading Levels," The Reading Teacher, XVII (January, 1964), 265-68.

Strang, Ruth. "Teaching Reading to the Culturally Disadvantaged in Secondary Schools," Journal of Reading, X (May, 1967), 527-35.

Weber, Rose-Marie. "The Study of Oral Reading Errors: A Survey of the Literature," Reading Research Quarterly, IV (Fall, 1968), 96-119.

E. ESSAYS AND ARTICLES IN COLLECTIONS

Carillo, Lawrence. "The Reading Program, K-6," SRA Reading Institute Extension Service, Unit One. Chicago: Science Research Associates, Inc., 1967.

Huss, Helen. "Basic Reading Skill Instruction in the Total Curriculum," Reading and the Elementary School Curriculum, David L. Shepherd, editor. Newark, Delaware: International Reading Association, 1969.

Ilg, Frances. "The Child from Three to Eight, With Implications for Reading," Issues and Innovations in the Teaching of Reading, Joe L. Frost, editor. Glenview, Illinois: Scott, Foresman and Company 1967. Pp. 81-88.

Skinner, B. F. "Teaching Machines," Readings for an Introduction to Psychology, Richard A. King, editor. New York: McGraw-Hill Book Company, Inc., 1961. Pp. 170-177.